ACADEMIC PRESS

J. Math. Anal. Appl. 270 (2002) 681–708

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

www.academicpress.com

Rectifiability results for singular and conjugate points of optimal exit time problems

Cristina Pignotti

Dipartimento di Matematica, Università di Roma "Tor Vergata," Via della Ricerca Scientifica, 00133 Rome, Italy

Received 22 February 2001

Submitted by U. Stadtmueller

Abstract

We consider the value function V of optimal control problems with exit time. Under suitable assumptions, through the study of the conjugate points, we prove that the closure of the singular set of V is rectifiable. Moreover, a sharper Hausdorff estimate is given on the set of the conjugate nonsingular points. © 2002 Elsevier Science (USA). All rights reserved.

0. Introduction

Optimal exit time problems play an important role in control theory (see, e.g., [17,20]). In these problems a closed set $\mathcal{K} \subset \mathbf{R}^n$, called the *target*, and a system

$$\begin{cases} \dot{y}(t) = f(y(t), u(t)), \\ y(0) = x \in \mathbf{R}^n \end{cases} \tag{1}$$

are given. Here f is a vector field and u is a *control*—i.e., a measurable function taking values in a given closed set $U \subset \mathbf{R}^m$ (the *control space*). We denote by

E-mail address: pignotti@axp.mat.uniroma2.it.

PII: S0022-247X(02)00110-5

$y^{x,u}(t)$ the solution of (1), and say that $y^{x,u}(t)$ is the trajectory starting at x with control u. We are interested in minimizing a *cost functional* of the form

$$J(x,u) = \int_0^{\tau(x,u)} L\big(y^{x,u}(t), u(t)\big)\,dt + g\big(y^{x,u}(\tau(x,u))\big),$$

where $\tau(x,u)$ denotes the first time at which the trajectory $y^{x,u}(\cdot)$ reaches the target $\mathcal{K}$. The most studied example (see, e.g., [8,11,12] and references therein) is the minimum time problem, which corresponds to the choice of $L \equiv 1$ and $g \equiv 0$.

The value function, defined as

$$V(x) = \inf_u J(x,u), \quad x \in \mathbf{R}^n,$$

can be used to derive optimality conditions on the trajectories of the system and, sometimes, an optimal feedback control. In addition, V is a viscosity solution—in the sense of Crandall and Lions [14,15]—of a suitable Hamilton–Jacobi equation. In general, V is not differentiable everywhere, and it is important to analyze the structure of its singularities and how these are related to the properties of the control system.

Some authors (see [7,22]) have studied the Lipschitz continuity of V, showing that it depends on suitable conditions on the final cost g and on the behaviour of the vector field f along the boundary of the target. If we call *singular* the points where V is not differentiable, and denote by Σ the set of all singular points, then the Lipschitz continuity of V implies, by Rademacher's theorem, that Σ has zero measure.

A finer regularity result has been obtained in [10], where it is proved that V is semiconcave, under some suitable additional assumptions on the smoothness of the data and of the boundary of the target. We recall that a function is called semiconcave if it can be locally represented as the sum of a concave function plus a smooth one. Therefore, semiconcave functions share many differentiability properties of concave functions. For instance, the singular set of a semiconcave function can be covered by countably many C^1—hypersurfaces of dimension $n-1$ (see, e.g., [3,25]). Sharp Hausdorff estimates from above and below for the singular set of a semiconcave function have been obtained in [1–3]. In [10] it is also shown that the set of all optimal trajectories starting at a given point x is related to the set of all limiting gradients of V at x. In particular, the singular points are exactly the initial conditions such that the control system has not a unique optimal trajectory.

In the present paper we prove some further regularity properties of the value function. In analogy with the problems in calculus of variations (see [19]) we introduce for our problem the set of *conjugate points*, which we denote by Γ. Then, we prove that $\overline{\Sigma} \subset \Sigma \cup \Gamma$ and that V is as smooth as the data in the complement of this set. This shows that it is interesting to bound from above the size of $\Sigma \cup \Gamma$.

One of the main results of this paper is that the set Γ of conjugate points is $\mathcal{H}^{n-1}$-rectifiable. Since Σ is also $\mathcal{H}^{n-1}$-rectifiable, the previous results imply that V is smooth in the complement of a closed set of codimension one. We point out that, for a general semiconcave function, the set of differentiability may even have empty interior.

We also obtain a sharper estimate on the set $\Gamma \setminus \Sigma$ of conjugate points which are not singular. In particular, we show that if the data of the control problem are of class C^∞, then the Hausdorff dimension of $\Gamma \setminus \Sigma$ is not greater than $n-2$. Roughly speaking, this shows that, even if the set Σ is not closed in general, the set $\overline{\Sigma} \setminus \Sigma$ is one dimension lower.

This analysis extends to optimal control problems with exit time the method applied in [9] for a problem of calculus of variations. A similar analysis has been done in [23] and in [24] for the distance function on a manifold and for solutions to Hamilton–Jacobi equations under suitable assumptions.

The outline of the paper is the following. In Section 2 we recall the definition of semiconcave functions and their basic differentiability properties. In Section 3 we give a precise formulation of our exit time problem and we recall the results about semiconcavity and optimality conditions obtained in [10]. In Section 4 we introduce the conjugate points and prove the properties recalled above about the closure of Σ. Moreover, we prove that $\Sigma \setminus \Gamma$ is contained in a locally finite union of smooth surfaces and we give a sufficient condition on a singular point to be conjugate. In Section 5 we prove the $\mathcal{H}^{n-1}$-rectifiability of the set $\overline{\Sigma}$ and, in Section 6, we prove the sharper estimate on the set $\Gamma \setminus \Sigma$ mentioned above. Finally, in Section 7 we give an equivalent characterization of conjugate points in the case of the minimum time function.

1. Notation and preliminaries

Given $r > 0$ and $x \in \mathbf{R}^n$, we set

$$B_r(x) = \{y \in \mathbf{R}^n\colon\ |x-y| < r\}$$

and we abbreviate $B_r = B_r(0)$. We denote by $p \cdot q$ or by $\langle p, q\rangle$ the usual scalar product of two vectors $p, q \in \mathbf{R}^n$. We recall that, if $v : \Omega \to \mathbf{R}$ and $x \in \Omega$, the *subdifferential* and the *superdifferential* of v at x are, respectively, the sets

$$D^- v(x) = \left\{p \in \mathbf{R}^n\colon\ \liminf_{y \to x} \frac{v(y) - v(x) - p \cdot (y-x)}{|y-x|} \geqslant 0\right\},$$
$$D^+ v(x) = \left\{p \in \mathbf{R}^n\colon\ \limsup_{y \to x} \frac{v(y) - v(x) - p \cdot (y-x)}{|y-x|} \leqslant 0\right\}.$$

If v is a locally Lipschitz function, then $D^- v(x)$ and $D^+ v(x)$ are compact convex sets. They are both nonempty if and only if v is differentiable at x and, in this case, they both contain only the gradient of v. For a locally Lipschitz function $v:\Omega\to\mathbf{R}$, we also define the set of *limiting gradients*

$$D^*v(x)=\Big\{p\colon\ \exists\{x_k\}\subset\Omega \text{ such that } x_k\to x,\\ v \text{ is differentiable at } x_k,\ p=\lim_{k\to\infty}Dv(x_k)\Big\},$$

where Dv denotes the usual gradient of v. This set is nonempty as a corollary of Rademacher's theorem. Now we recall the definition of semiconcavity.

Definition 1.1. A continuous function $v:\Omega\to\mathbf{R}$, with $\Omega\subset\mathbf{R}^n$, is called *semiconcave* if, for any convex $K\subset\subset\Omega$, there exists $c_K>0$ such that

$$v(x_1)+v(x_2)-2v\left(\frac{x_1+x_2}{2}\right)\leqslant c_K|x_1-x_2|^2, \tag{2}$$

for any $x_1,x_2\in K$.

It is easy to see that condition (2) holds if and only if $v(x)-2c_K|x|^2$ is concave in K. Therefore, semiconcave functions retain some differentiability properties of concave functions. For example, they are twice differentiable almost everywhere. There is a more general definition of semiconcavity (see [3]) which will not be used in this paper. The superdifferential of a semiconcave function enjoys the following properties (see [13]).

Theorem 1.1. *Let $v:\Omega\to\mathbf{R}$ be semiconcave. Then v is locally Lipschitz in Ω, and*

$$D^+v(x)=\operatorname{co}D^*v(x),\quad \forall x\in\Omega, \tag{3}$$

where $\operatorname{co}(\cdot)$ denotes the convex hull. Therefore the superdifferential D^+v is nonempty at each point. Moreover, D^+v is upper semicontinuous; that is,

$$\textit{if}\quad x_k\to x,\quad p_k\in D^+v(x_k),\quad p_k\to p,\quad \textit{then}\quad p\in D^+v(x).$$

In particular, if Dv exists everywhere in Ω, then $v\in C^1(\Omega)$.

We recall that, given a set $A\subset\mathbf{R}^n$ and a real number $\alpha\in[0,n]$, the *α-dimensional Hausdorff measure* of A is defined by

$$\mathcal{H}^\alpha(A)=\frac{\omega_\alpha}{2^\alpha}\sup_{\delta>0}\inf\left\{\sum_{i=1}^\infty(\operatorname{diam}B_i)^\alpha\colon\ A\subset\bigcup_{i=1}^\infty B_i,\ \operatorname{diam}B_i<\delta\right\},$$

where $\omega_0=1$, ω_α is the Lebesgue measure of the unit ball in $\mathbf{R}^\alpha$ if $\alpha\geqslant1$ is an integer, and ω_α is a suitable positive constant otherwise. The *Hausdorff dimension* of A is given by

$$\mathcal{H}-\dim(A)=\inf\big\{\alpha>0\colon\ \mathcal{H}^\alpha(A)=0\big\}.$$

Let $k \in \{0, \dots, n\}$. We say that a set $C \subset \mathbf{R}^n$ is (countably) *$\mathcal{H}^k$-rectifiable* if it can be covered, up to a $\mathcal{H}^k$ negligible set, by the union of a countable family of C^1-hypersurfaces of dimension k.

Definition 1.2. Given a semiconcave function $v : \Omega \to \mathbf{R}$, the set

$$\Sigma(v) = \bigl\{x \in \Omega:\ \nexists Dv(x)\bigr\}$$

is called the *singular set* of v.

The next theorem is essentially due to Zajíček (see [25]). Related and improved versions of this result can be found in [3] and [1].

Theorem 1.2. *If $v : \Omega \subset \mathbf{R}^n \to \mathbf{R}$ is semiconcave, then the singular set $\Sigma(v)$ is $\mathcal{H}^{n-1}$-rectifiable.*

2. The optimal exit time problem

Suppose that a closed set with compact boundary $\mathcal{K} \subset \mathbf{R}^n$ (called the *target*), a compact set $U \subset \mathbf{R}^m$ and a continuous function $f : \mathbf{R}^n \times U \to \mathbf{R}^n$ are given. We fix $x \in \mathbf{R}^n$ and a measurable function $u : \mathbf{R}_+ \to U$ and we consider the system

$$\begin{cases} \dot{y}(t) = f(y(t), u(t)), \\ y(0) = x. \end{cases} \tag{4}$$

The function u is called a *control* for (4). In the following suitable assumptions will be made on f to ensure the global existence of a unique solution to (4), which will be denoted by $y^{x,u}(\cdot)$. We then set

$$\tau(x, u) = \min\bigl\{t \geqslant 0:\ y^{x,u}(t) \in \mathcal{K}\bigr\},$$

with the convention that $\tau(x, u) = +\infty$ if $y^{x,u}(t) \notin \mathcal{K}$ for all $t \geqslant 0$. In other words, $\tau(x, u)$ is the time at which $y^{x,u}$ reaches the target; we call it the *exit time* of the trajectory. For simplicity of notation we set

$$y^{x,u}_\tau := y^{x,u}(\tau(x, u))$$

to denote the point where the trajectory reaches the target. We denote by $\mathcal{R}$ the set of all x such that $\tau(x, u) < +\infty$ for some control u and call $\mathcal{R}$ the *attainable set*.

We then suppose that two continuous functions $L : \mathbf{R}^n \times U \to \mathbf{R}$ (called *running cost*) and $g : \mathbf{R}^n \to \mathbf{R}$ (the *terminal cost*) are given, with $L \geqslant 0$, and we consider the functional

$$J(x, u) = \int_0^{\tau(x,u)} L\bigl(y^{x,u}(s), u(s)\bigr)\, ds + g\bigl(y^{x,u}_\tau\bigr).$$

We are interested, for $x \in \mathcal{R}$, in minimizing $J(x,u)$ over all measurable controls $u:[0,\infty) \to U$. A control $\bar{u}$ and the corresponding trajectory $y^{x,\bar{u}}$ are called *optimal* for the point x if

$$J(x,\bar{u}) = \min_u J(x,u).$$

The *value function* of this problem is defined as

$$V(x) = \inf\bigl\{J(x,u):\ u:[0,+\infty[\to U\bigr\}, \quad x \in \mathcal{R}. \tag{5}$$

From the definition of V one can easily obtain the dynamic programming principle: for any $x \in \mathbf{R}^n$ and $u:[0,\infty[\to U$,

$$V(x) \leqslant \int_0^t L\bigl(y^{x,u}(s), u(s)\bigr)\,ds + V\bigl(y^{x,u}(t)\bigr), \quad \forall t \in [0, \tau(x,u)], \tag{6}$$

where the equality holds if u is optimal.

The value function V is a viscosity solution of the corresponding *Hamilton–Jacobi–Bellman equation*, which, for this problem, has the form

$$H\bigl(x, DV(x)\bigr) = 0, \tag{7}$$

where

$$H(x,p) = \max_{u\in U}\bigl[-f(x,u)\cdot p - L(x,u)\bigr]. \tag{8}$$

We recall that the *minimum time optimal control* problem for system (4) with target $\mathcal{K}$ consists of minimizing $\tau(x,u)$ over all measurable controls $u:\mathbf{R}_+ \to U$. It is a particular exit time problem, corresponding to the choices $L \equiv 1$ and $g \equiv 0$. The value function for this problem reduces to

$$T(x) = \inf\bigl\{\tau(x,u):\ u:[0,+\infty[\to U\bigr\}, \quad x \in \mathcal{R}, \tag{9}$$

and is called *minimum time function*.

In this paper, fixed $k \in \mathbf{N}$, $k \geqslant 2$, we assume the following hypotheses:

(A1) For all $x \in \mathbf{R}^n$ the set

$$F(x) := \bigl\{(v,\lambda) \in \mathbf{R}^{n+1}:\ \exists u \in U \text{ such that } v = f(x,u),\ \lambda \geqslant L(x,u)\bigr\}$$

is convex.

(A2) The functions f, L are of class C^{k+1} in both arguments and the boundary $\partial\mathcal{K}$ of the target set is a $(n-1)$-manifold of class C^{k+1}. Moreover, g is of class C^{k+1} in a neighbourhood $\mathcal{N}$ of $\partial\mathcal{K}$.

(A3) There exist constants N, $\alpha > 0$ such that $|f(x,u)| \leqslant N$ and $L(x,u) \geqslant \alpha$ $\forall x \in \mathbf{R}^n$ and $\forall u \in U$. Moreover, denoting by G the Lipschitz constant of g in $\mathcal{N}$, we assume

$$G < \alpha/N. \tag{10}$$

(A4) There exists $\gamma > 0$ such that, for all $x \in \partial\mathcal{K}$, denoted by $\nu(x)$ the outer normal unit vector to $\mathcal{K}$ at x,

$$\min_{u\in U} \nu(x)\cdot f(x,u) \leqslant -\gamma.$$

Moreover, we make the following assumptions:

(H1) For any $x \in \mathbf{R}^n$, if $H(x,p)=0$ for all p in a convex set C, then C is a singleton.

(H2) For all $(x,p) \in \mathbf{R}^n \times (\mathbf{R}^n \setminus \{0\})$ there exists a unique $u^* \in U$ such that

$$-f(x,u^*)\cdot p - L(x,u^*) = \max_{u\in U}\bigl[-f(x,u)\cdot p - L(x,u)\bigr], \tag{11}$$

and the function $u^*\colon (x,p) \to u^*(x,p)$ is of class C^k in $\mathbf{R}^n \times (\mathbf{R}^n \setminus \{0\})$.

Remark 2.1. Actually, we only need that u^* is of class C^k in an open neighbourhood of the set $\{(x,p) \in \mathbf{R}^n \times (\mathbf{R}^n \setminus \{0\})\colon H(x,p)=0\}$. This will be clear from our arguments.

Assumption (A1) is a condition ensuring the existence of optimal trajectories. In fact, one can prove, by standard techniques, the following result.

Theorem 2.1. *Under hypotheses* (A1)–(A3)*, there exists an optimal control for any choice of initial point $x \in \mathcal{R}$. Moreover, the uniform limit of optimal trajectories is an optimal trajectory*; *that is, if y_k are trajectories converging uniformly to y and every y_k is optimal for the point $x_k := y_k(0)$, then y is optimal for $x := \lim x_k$.*

Remark 2.2. The restriction (10) in assumption (A3) can be regarded as a compatibility condition on the terminal cost g. Together with our other assumptions, it ensures the continuity of the value function (see Proposition IV.3.7 in [7]).

In the following we denote by A^T the transpose of a given matrix A.

Remark 2.3. From assumptions (A2) and (H2), it follows that the Hamiltonian H is of class C^{k+1} in $\mathbf{R}^n \times (\mathbf{R}^n \setminus \{0\})$. In fact, in [10], using standard arguments we proved that for any $(x,p) \in \mathbf{R}^n \times (\mathbf{R}^n \setminus \{0\})$,

$$\begin{aligned} H_p(x,p) &= -f\bigl(x,u^*(x,p)\bigr),\\ H_x(x,p) &= -f_x^T\bigl(x,u^*(x,p)\bigr)p - L_x\bigl(x,u^*(x,p)\bigr). \end{aligned} \tag{12}$$

Now we give some examples for which hypothesis (H2) holds.

Example 2.1. Let $U = \overline{B}_r \subset \mathbf{R}^n$ and $f(x,u) = \sigma(x)u$, where $\sigma(x)$ is a $n \times n$ nonsingular matrix for $x \in \mathbf{R}^n$. Furthermore, we assume that the running cost L depends only on the state variable; that is, $L(x,u) = L(x)$. In this case it is easy to see that

$$u^*(x,p) = -r\frac{\sigma^T(x)p}{|\sigma^T(x)p|}$$

verifies the relation (11). Moreover, u^* has the same x-regularity as $\sigma(\cdot)$ for $p \neq 0$. Therefore, if $\sigma(\cdot)$ is of class C^k, then assumption (H2) holds.

Remark 2.4. More generally, in the case that the running cost L depends only on the state, if we assume that $\partial(f(x,U))$ is a C^{k+1} manifold with positive curvature and that $f_u(x,u)$ is nonsingular for every x, then hypothesis (H2) holds. We refer to [12] for the proof of such a result in the case of the minimum time function. Obviously, the proof remains valid for exit time problems with $L(x,u) = L(x)$.

Now we give an example with L depending also on the control variables.

Example 2.2. Consider the dynamics $f(x,u) = \sigma(x)u$, where $\sigma(x)$ is a $n \times n$ nonsingular matrix, $x \in \mathbf{R}^n$. We assume that for a positive number N,

$$|w| \leqslant |\sigma(x)w| \leqslant N|w|, \quad \forall w \in \mathbf{R}^n, \ \forall x \in \mathbf{R}^n. \tag{13}$$

Consider the running cost

$$L(x,u) = l(x) + \frac{1}{2}|u|^2, \quad l(x) > \alpha > 0,$$

and the control space $U = \overline{B}_r$, where $r < \min\{1, \alpha/N\}$. We can easily compute that

$$u^*(x,p) = \begin{cases} -\sigma^T(x)p, & |\sigma^T(x)p| \leqslant r, \\ -r\frac{\sigma^T(x)p}{|\sigma^T(x)p|}, & |\sigma^T(x)p| > r. \end{cases} \tag{14}$$

The function u^* is locally Lipschitz continuous in $\mathbf{R}^n \times (\mathbf{R}^n \setminus \{0\})$ but it is not of class C^k.

We consider the set $W = \{(x,p)\colon H(x,p) = 0\}$. For $(x,p) \in W$, we have

$$-\sigma(x)u^*(x,p) \cdot p - l(x) - \frac{1}{2}\left|u^*(x,p)\right|^2 = 0.$$

Then, by our assumptions on σ, l, r, we obtain $|p| > r$ from which follows, using (13), $|\sigma^T(x)p| > r$. Therefore, for all (x,p) in an open neighbourhood of W,

$$u^*(x,p) = -r\frac{\sigma^T(x)p}{|\sigma^T(x)p|}.$$

So, if $\sigma \in C^k$, then u^* is of class C^k in a neighbourhood of the 0-level set W of the Hamiltonian. This is sufficient for our purposes (see Remark 2.1).

We have the following result, proved in [10].

Theorem 2.2. *Under hypotheses* (A1)–(A4) *the value function V is semiconcave in $\mathcal{R} \setminus \mathcal{K}$.*

Remark 2.5. Actually, in [10], the above result is proved under weaker assumptions on f, L, g and $\mathcal{K}$. In particular, $\partial\mathcal{K}$ is only assumed to satisfy an interior sphere condition, while f, L, g are just supposed to be semiconcave in the first argument. The semiconcavity in the particular case of the minimum time function has been proved in [11].

We recall some other results obtained in [10] that will be useful in the following.

Proposition 2.1. *Let $y(\cdot)$ be an optimal trajectory with exit time τ. Then V is differentiable at $y(t)$ for all $t \in (0, \tau)$.*

For the control system considered we can give a maximum principle in the following form.

Proposition 2.2. *Assume properties* (A2)–(A4). *Let $x \in \mathcal{R} \setminus \mathcal{K}$ and let $\bar{u}$ be an optimal control for x. Set for simplicity*

$$y(t) := y^{x,\bar{u}}(t), \qquad \tau := \tau(x, \bar{u}), \qquad \xi := y_\tau^{x,\bar{u}},$$

and denote by $\nu(\xi)$ the outer normal to $\mathcal{K}$ at ξ. Then, there exists a unique positive number $\mu(\xi)$ such that $H(\xi, Dg(\xi) + \mu(\xi)\nu(\xi)) = 0$. Moreover, if $p : [0, \tau] \to \mathbf{R}^n$ is the solution of the system

$$\begin{cases} \dot{p}(t) = -f_x^T(y(t), \bar{u}(t))p(t) - L_x(y(t), \bar{u}(t)), \\ p(\tau) = Dg(\xi) + \mu(\xi)\nu(\xi), \end{cases} \tag{15}$$

where f_x^T denotes the transpose of the matrix f_x, then p satisfies, for a.e. $t \in [0, \tau]$ and for all $u \in U$,

$$-p(t) \cdot f\big(y(t), \bar{u}(t)\big) - L\big(y(t), \bar{u}(t)\big) \geqslant -p(t) \cdot f\big(y(t), u\big) - L\big(y(t), u\big). \tag{16}$$

Given an optimal trajectory y, we say that p is the *dual arc* associated to y if it satisfies the properties of Proposition 2.2.

Proposition 2.3. *Let y be an optimal trajectory with exit time τ, and let p be its dual arc. Then, $p(t) \neq 0$, $\forall t \in [0, \tau]$. Moreover,*

$$p(t) = DV(y(t)), \quad \forall t \in (0, \tau).$$

Remark 2.6. In particular, by Propositions 2.2 and 2.1, along any pair optimal trajectory–dual arc associated $(y(\cdot), p(\cdot))$ we have $H(y(t), p(t)) = 0$.

Theorem 2.3. *Assume properties* (A1)–(A4) *and* (H1), (H2). *Let* $x \in \mathcal{R} \setminus \mathcal{K}$ *and* $q \in D^*V(x)$. *Consider the solution* (y, p) *of*

$$\begin{cases} \dot{y}(t) = -H_p(y(t), p(t)), \\ \dot{p}(t) = H_x(y(t), p(t)), \\ y(0) = x, \\ p(0) = q. \end{cases} \tag{17}$$

Then y *is an optimal trajectory for* x *and* p *is the dual arc associated to* y. *Conversely, any optimal trajectory for* x *can be obtained solving* (17) *for a suitable choice of* $q \in D^*V(x)$.

Remark 2.7. Actually, Proposition 2.2 and Theorem 2.3 are proved in [10] under weaker assumptions. Here, for the additional regularity of the Hamiltonian H, from Theorem 2.3 it follows that y and p are of class C^{k+1}.

From Theorems 1.1 and 2.2 we can deduce the following result.

Corollary 2.1. *Under the assumptions of Theorem* 2.3, V *is differentiable at a point* $x \in \mathcal{R} \setminus \mathcal{K}$ *if and only if the optimal trajectory for* x *is unique.*

3. The backward Hamiltonian system

Given $\xi \in \partial\mathcal{K}$, we call $\nu(\xi)$ the normal to $\partial\mathcal{K}$ at ξ and denote by $(Y(\xi, \cdot), P(\xi, \cdot))$ the solution of the backward Hamiltonian system

$$\begin{cases} \dot{Y}(t) = H_p(Y(t), P(t)), \\ \dot{P}(t) = -H_x(Y(t), P(t)), \end{cases} \tag{18}$$

with initial conditions

$$\begin{cases} Y(0) = \xi, \\ P(0) = Dg(\xi) + \mu(\xi)\nu(\xi), \end{cases} \tag{19}$$

where $\mu(\xi) > 0$ is the unique positive constant such that

$$H\big(\xi, Dg(\xi) + \mu(\xi)\nu(\xi)\big) = 0. \tag{20}$$

We recall that the Hamiltonian H is smooth only on $\mathbf{R}^n \times (\mathbf{R}^n \setminus \{0\})$, but it is not difficult to prove that the solutions of system (18) with initial conditions (19) are defined for any positive time. First of all, we note that $P(\xi, 0) \neq 0$. Indeed, if $P(\xi, 0) = 0$, we should have

$$H\big(\xi, Dg(\xi) + \mu(\xi)\nu(\xi)\big) = \sup_{u \in U}\big\{-L(\xi, u)\big\} \leqslant -\alpha,$$

in contradiction with (20). So, for any $\xi \in \partial\mathcal{K}$ there exists a local solution of system (18) with conditions (19). Moreover, we can give the following lemma.

Lemma 3.1. *Let* $(Y(\xi,\cdot), P(\xi,\cdot))$ *be the maximal solution of system* (18) *with initial conditions* (19). *Then* $(Y(\xi,\cdot), P(\xi,\cdot))$ *is defined for all* $t \in [0,+\infty)$.

Proof. We argue by contradiction. We suppose that there exists a finite time t^* such that system (18), with conditions (19), can be defined only on the interval $[0,t^*)$. Recalling Remark 2.3 and the regularity assumptions on f and L, the solution (Y,P) cannot blow up in a finite time. Then, we have $P(\xi,t^*)=0$. It is easy to see that the Hamiltonian H is constant along the pair $(Y(\xi,\cdot), P(\xi,\cdot))$ solution of (18). So,

$$H\big(Y(\xi,t), P(\xi,t)\big) = 0, \quad t \in [0,t^*),$$

that implies $P(\xi,t) \neq 0$ and, more precisely,

$$|P(\xi,t)| \geqslant \frac{\alpha}{N}, \quad t \in [0,t^*).$$

Then, by continuity, $P(\xi,t^*) \neq 0$ and this contradicts our previous assumption. Therefore $(Y(\xi,\cdot), P(\xi,\cdot))$ is defined for any $t \in [0,+\infty)$. □

Definition 3.1. We say that a characteristic $Y(\xi,\cdot)$ is *optimal* in some interval $[0,\tau]$ if it coincides with an optimal trajectory run backward in time. In this case,

$$Y(\xi,t) = y^{x,u}(\tau - t), \quad t \in [0,\tau],$$

for some $x \in \mathcal{R} \setminus \mathcal{K}$ and u optimal control for x.

Proposition 3.1. *For any* $\xi \in \partial\mathcal{K}$, *we have*

$$H_p\big(\xi, Dg(\xi) + \mu(\xi)\nu(\xi)\big) \cdot \nu(\xi) \neq 0. \tag{21}$$

Proof. By (12), we have to prove that $f(\xi,u^*) \cdot \nu(\xi) \neq 0$, where, for simplicity, $u^* = u^*(\xi, Dg(\xi) + \mu(\xi)\nu(\xi))$. By definition of u^*,

$$\begin{aligned}
0 &= -f(\xi,u^*) \cdot \big(Dg(\xi) + \mu(\xi)\nu(\xi)\big) - L(\xi,u^*) \\
&\leqslant NG - \alpha - \mu(\xi) f(\xi,u^*) \cdot \nu(\xi) < -\mu(\xi) f(\xi,u^*) \cdot \nu(\xi),
\end{aligned}$$

where we used (10). Therefore, being $\mu(\xi) > 0$, $f(\xi,u^*) \cdot \nu(\xi) < 0$, $\forall \xi \in \partial\mathcal{K}$. □

Remark 3.1. Geometrically, relation (21) implies that the optimal trajectories cannot arrive tangentially at the target.

By (21) the surface $\partial\mathcal{K}$ is noncharacteristic for the data g. So, there exists a time $t^* > 0$ such that $Y(\cdot,\cdot)$ is a diffeomorphism in $\partial\mathcal{K} \times [0,t^*)$ (see, e.g., [16]). Therefore, by standard techniques, any characteristic $Y(\xi,\cdot)$ coincides with an optimal trajectory run backward in time, at least for small times.

As an easy consequence of the previous proposition, we have the following result.

Proposition 3.2. *The function* $\mu : \partial\mathcal{K} \to \mathbf{R}^+$ *is of class* C^k.

Proof. We observe that (21) can be rewritten as

$$D_\lambda H\big(\xi, Dg(\xi) + \lambda\nu(\xi)\big) \neq 0.$$

Then, by the implicit function theorem, $\lambda = \mu(\xi)$ is k-times continuously differentiable at any $\xi \in \partial\mathcal{K}$. □

We can give an immediate corollary of Proposition 3.2.

Corollary 3.1. *The functions* Y *and* P *are of class* C^k *on* $\partial\mathcal{K} \times [0, \infty)$.

For $x \in \mathcal{R}$ we set

$$\mathcal{F}(x) = \big\{\xi \in \mathcal{K}\colon\ \exists y^{x,u}(\cdot) \text{ optimal for } x,\ y^{x,u}(\tau(x,u)) = \xi\big\}.$$

By Theorem 2.3 we deduce that for any $x \in \mathcal{R}$

$$\xi \in \mathcal{F}(x) \quad \Rightarrow \quad Y\big(\xi, \tau(x,u)\big) = x, \tag{22}$$

with u optimal control for x such that $y^{x,u}(\tau(x,u)) = \xi$.

By differentiating system (18) with respect to time we obtain

$$\begin{cases} \dot{Y}_t = H_{px}(Y,P)Y_t + H_{pp}(Y,P)P_t, \\ \dot{P}_t = -H_{xx}(Y,P)Y_t - H_{xp}(Y,P)P_t, \end{cases} \tag{23}$$

and differentiating with respect to ξ we obtain, formally,

$$\begin{cases} \dot{Y}_\xi = H_{px}(Y,P)Y_\xi + H_{pp}(Y,P)P_\xi, \\ \dot{P}_\xi = -H_{xx}(Y,P)Y_\xi - H_{xp}(Y,P)P_\xi. \end{cases} \tag{24}$$

Therefore,

$$\begin{cases} \dot{Y}_{\xi,t} = H_{px}(Y,P)Y_{\xi,t} + H_{pp}(Y,P)P_{\xi,t}, \\ \dot{P}_{\xi,t} = -H_{xx}(Y,P)Y_{\xi,t} - H_{xp}(Y,P)P_{\xi,t}, \end{cases} \tag{25}$$

where $Y_{\xi,t}(\cdot)$ and $P_{\xi,t}(\cdot)$ denote respectively the Jacobian of Y and P with respect to the pair (ξ, t).

Remark 3.2. Fixed $\xi_0 \in \partial\mathcal{K}$, $t_0 > 0$, we need to explain the sense in which $Y_{\xi,t}(\xi_0, t_0)$ and $P_{\xi,t}(\xi_0, t_0)$ have to be understood. Since $\partial\mathcal{K}$ is a smooth $(n-1)$-dimensional manifold, for any $\xi_0 \in \partial\mathcal{K}$ there exist an open neighbourhood $I_{\xi_0} \subset \partial\mathcal{K}$ and a parametrization ψ of class C^{k+1}, with inverse φ of class C^{k+1},

$$\psi : I_{\xi_0} \subset \partial\mathcal{K} \to \psi(I_{\xi_0}) \subset \mathbf{R}^{n-1}, \quad \xi \to \eta,$$

where $\psi(I_{\xi_0})$ is an open neighbourhood of $\eta_0 = \psi(\xi_0)$. Therefore, with $Y_{\xi,t}(\xi_0, t_0)$ and $P_{\xi,t}(\xi_0, t_0)$ we denote the Jacobians of $Y(\varphi(\cdot), \cdot)$ and $P(\varphi(\cdot), \cdot)$, evaluated at the point (η_0, t_0), with respect to the time t and the coordinates $\eta \in \mathbf{R}^{n-1}$,

$$Y_{\eta,t}\bigl(\varphi(\eta_0), t_0\bigr), \; P_{\eta,t}\bigl(\varphi(\eta_0), t_0\bigr).$$

We note that this definition is dependent on the particular parametrization chosen, but for our purposes this is not important. Indeed, to obtain our results, we are only interested in the set where $\det Y_{\xi,t}$ is zero or in the sets where $Y_{\xi,t}$, $P_{\xi,t}$ have a given rank and these sets are independent of the choice of parametrization.

Proposition 3.3. *Let $\{x_j\}_{j\in\mathbf{N}}$ be a sequence in $\mathcal{R} \setminus \mathcal{K}$ convergent to some $x \in \mathcal{R} \setminus \mathcal{K}$. For any $j \in \mathbf{N}$ consider a point $\xi_j \in \mathcal{F}(x_j)$ and let τ_j be the exit time of the optimal trajectory starting at x_j and arriving in ξ_j.*

Then, up to a subsequence, there exist $\xi \in \partial\mathcal{K}$, $\tau \in \mathbf{R}^+$, such that

$$\xi_j \to \xi, \quad \tau_j \to \tau, \quad \text{as } j \to \infty,$$

with $Y(\xi, \tau) = x$. Moreover, the characteristic $Y(\xi, t)$ is optimal in $[0, \tau]$.

Proof. Since $\partial\mathcal{K}$ is bounded, up to a subsequence, $\{\xi_j\}_{j\in\mathbf{N}}$ converges to ξ. Now we prove that $\{\tau_j\}_{j\in\mathbf{N}}$ is bounded. Observing that the value function V is continuous, we have

$$V(x_j) \leqslant |V(x)| + c_1, \quad \forall j \in \mathbf{N}, \tag{26}$$

for some positive constant c_1. Moreover, since the terminal cost g is continuous we have

$$g(\xi_j) \geqslant -c_2 + |g(\xi)|, \quad \forall j \in \mathbf{N}, \tag{27}$$

for a positive constant c_2. Now, we can argue by contradiction. Suppose that the sequence $\{\tau_j\}_{j\in\mathbf{N}}$ is unbounded. This means that for any $M > 0$, there exists $j(M) \in \mathbf{N}$ such that $\tau_{j(M)} > M$. Then,

$$V(x_{j(M)}) = \int_0^{\tau_{j(M)}} L\bigl(y_{j(M)}(t), u_{j(M)}(t)\bigr)\, dt + g(\xi_{j(M)}),$$

where we denoted with $(y_{j(M)}, u_{j(M)})$ the pair optimal trajectory–control for $x_{j(M)}$. From (27), recalling hypothesis (A2), it follows that

$$V(x_{j(M)}) \geqslant M\alpha - c_2 + |g(\xi)|. \tag{28}$$

Since the right-hand side in (28) goes to infinity as $M \to \infty$, we have a contradiction with (27). Thus, the sequence $\{\tau_j\}_{j\in\mathbf{N}}$ is bounded in $(0, +\infty)$. Possibly considering a subsequence, we can suppose that $\{\tau_j\}_{j\in\mathbf{N}}$ converges to some time τ. By the uniqueness of the limit,

$$x_j = Y(\xi_j, \tau_j) \to Y(\xi, \tau) = x, \quad \text{as } j \to \infty.$$

Recalling Corollary 3.1 and Theorem 2.1, we can conclude that $Y(\xi, t)$, for $t \in [0, \tau]$, is an optimal trajectory run backward in time. $\square$

Remark 3.3. In particular, by the previous proposition, it follows that the set-valued map $\mathcal{F}(\cdot)$ is upper semicontinuous.

Now, we introduce some definitions.

Definition 3.2. A point $x_0 \in \mathcal{R} \setminus \mathcal{K}$ is called *regular* if $\mathcal{F}(x_0)$ is a singleton. A point $x_0 \in \mathcal{R} \setminus \mathcal{K}$ is called *conjugate* if there exists $\xi_0 \in \mathcal{F}(x_0)$, $\xi_0 = y^{x_0,u}(\tau(x_0, u))$, such that

$$\det Y_{\xi,t}\big(\xi_0, \tau(x_0, u)\big) = 0.$$

We denote by Γ the set of the conjugate points of V.

Observe that, by the maximum principle, we cannot have two optimal trajectories starting at the point x_0 and arriving in the same $\xi \in \partial\mathcal{K}$. In fact, if there were two optimal trajectories y_1, y_2, then the pairs optimal trajectory–dual arc (y_1, p_1), (y_2, p_2) would satisfy the same Hamiltonian system with the same final conditions. So, by Corollary 2.1, the set of regular points coincides with the set where V is differentiable. If x_0 is a regular point we denote, for simplicity, $\tau(x_0)$ the exit time of the unique optimal trajectory starting at x_0.

We have the following result.

Theorem 3.1. *Let us denote with A the set of all the regular points $x \in \mathcal{R} \setminus \mathcal{K}$ that are not conjugate. Then A is open and $V \in C^k(A)$.*

Proof. Suppose that the point $x_0 \in A$ is the limit of a sequence $\{x_j\}_{j\in\mathbf{N}}$ in $(\mathcal{R} \setminus \mathcal{K}) \setminus A$. Possibly considering a subsequence we can suppose that the points x_j are either all singular or all regular and conjugate.

First we consider the case that x_j are singular. By (22), for any j there exist two different $\xi_j^1, \xi_j^2 \in \partial\mathcal{K}$ such that

$$x_j = Y\big(\xi_j^1, \tau_j^1\big) = Y\big(\xi_j^2, \tau_j^2\big), \tag{29}$$

where τ_j^1 and τ_j^2 are the exit times of optimal trajectories starting at x_j and arriving respectively in ξ_j^1 and ξ_j^2. Again extracting a subsequence we may assume that the sequences $\{\xi_j^1\}_{j\in\mathbf{N}}$ and $\{\xi_j^2\}_{j\in\mathbf{N}}$ converge to some limit ξ^1 and ξ^2, respectively. Then, by Proposition 3.3,

$$x_0 = Y(\xi^1, \tau^1) = Y(\xi^2, \tau^2),$$

where τ^1 and τ^2 are the exit times of optimal trajectories starting at x_0 and arriving respectively in ξ^1 and ξ^2. Since x_0 is a regular point this implies that $\tau^1 =$

$\tau^2 = \tau(x_0)$ and $\xi^1 = \xi^2$. Moreover, since x_0 is not conjugate, $Y(\cdot\,,\cdot)$ is a one-to-one correspondence in a neighbourhood of $(\xi^1, \tau(x_0))$. But this contradicts (29).

Now we consider the case when the points x_j are all regular and conjugate. Then, for any x_j there exist $\xi_j \in \partial\mathcal{K}$ and an optimal trajectory with exit time τ_j such that

$$x_j = Y(\xi_j, \tau_j).$$

Moreover, the Jacobian of Y at (ξ_j, τ_j) is singular. Up to a subsequence, we may assume that $\{\xi_j\}_{j\in\mathbf{N}}$ converges to some $\xi_0 \in \partial\mathcal{K}$, $\{\tau_j\}_{j\in\mathbf{N}}$ converges to some τ. Therefore,

$$Y(\xi_0, \tau) = \lim_{j\to\infty} Y(\xi_j, \tau_j) = x_0$$

and

$$\det Y_{\xi,t}(\xi_0, \tau) = \lim_{j\to\infty} \det Y_{\xi,t}(\xi_j, \tau_j) = 0.$$

This implies that $\xi_0 \in \mathcal{F}(x_0)$ and that x_0 is conjugate, in contradiction with our assumption. Then A is open.

Now, let $x_0 \in A$. Since A is open, $\mathcal{F}(\cdot)$ is single-valued and continuous in a neighbourhood of x_0. In addition,

$$Y\big(\mathcal{F}(x), \tau(x)\big) = x$$

for any x near x_0. It follows that $(\mathcal{F}, \tau)$ is the local inverse of Y near x_0. Since x_0 is not conjugate, the Jacobian of Y at $(\mathcal{F}(x_0), \tau(x_0))$ is nonsingular and then τ and $\mathcal{F}$ are as smooth as Y in a neighbourhood of x_0. So, in this neighbourhood of x_0,

$$\begin{aligned} &V\big(Y\big(\mathcal{F}(x), \tau(x)\big)\big) \\ &\quad = \int_0^{\tau(x)} L\Big(Y\big(\mathcal{F}(x), \tau(x) - s\big), \\ &\qquad\qquad u^*\big(Y\big(\mathcal{F}(x), \tau(x) - s\big), P\big(\mathcal{F}(x), \tau(x) - s\big)\big)\Big)\, ds \\ &\qquad + g(\mathcal{F}(x)). \end{aligned}$$

Then, recalling our regularity assumptions on L, g, u^*, Corollary 3.1 and the above remark on $\mathcal{F}$ and τ, it is easy to see that the value function V is in $C^k(A)$. $\square$

Remark 3.4. Note that, by the proof of the above theorem, the function that associates to any regular point x its exit time $\tau(x)$ belongs to $C^k(A)$.

As an immediate consequence of Theorem 3.1 we have the following result.

Corollary 3.2. *Let $\overline{\Sigma}$ be the closure in $\mathcal{R}\setminus\mathcal{K}$ of the set of the singular points of V. Then,*

$$\overline{\Sigma} \subset \Sigma \cup \Gamma.$$

We now give a sufficient condition for a point to be conjugate in terms of the set of limiting gradients.

Theorem 3.2. *Let $x_0 \in \mathcal{R}\setminus\mathcal{K}$. If $D^*V(x_0)$ is an infinite set then $x_0 \in \Gamma$.*

Proof. Let us assume that x_0 is not conjugate. We prove that the points of $\mathcal{F}(x_0)$ are isolated.

In fact, if $\xi_0 \in \mathcal{F}(x_0)$, then $Y_{\xi,t}(\xi_0,\tau_0)$ is nonsingular, where τ_0 is the exit time of the optimal trajectory starting at x_0 and arriving in ξ_0. And so

$$Y(\xi_0,\tau_0) \neq Y(\xi,t), \quad \text{for } (\xi,t) \in I_{\xi_0}\times J_{\tau_0}, \tag{30}$$

where $I_{\xi_0} \subset \partial\mathcal{K}$ is an open neighbourhood of ξ_0 and $J_{\tau_0} \subset (0,+\infty)$ is an open neighbourhood of τ_0. If we take I_{ξ_0} sufficiently small we prove that

$$I_{\xi_0} \cap \mathcal{F}(x_0) = \{\xi_0\}. \tag{31}$$

We can argue by contradiction. Let us suppose that for all neighbourhoods $I^j_{\xi_0}$ of ξ_0 of type

$$I^j_{\xi_0} = B(\xi_0, 1/j) \cap \partial\mathcal{K}$$

there exists $\xi_j \in I^j_{\xi_0}$ such that $\xi_j \in \mathcal{F}(x_0)$. Then, $Y(\xi_j,\tau_j) = x_0$ for a positive time τ_j. Recalling (30), we have $\tau_j \notin J_{\tau_0}$. From Proposition 3.3, we have that, up to a subsequence,

$$(\xi_j,\tau_j) \to (\xi_0,\tau_0), \quad \text{as } j \to \infty,$$

with $Y(\xi_0,\tau_0) = x_0$, in contradiction with the fact that $\tau_j \notin J_{\tau_0}$, $j \in \mathbf{N}$. Therefore, there exists a neighbourhood I_{ξ_0} of ξ_0 such that (31) holds.

Since $\mathcal{F}(x_0)$ is bounded, this implies that $\mathcal{F}(x_0)$ is finite. So, by Theorem 2.3, also $D^*V(x_0)$ is finite, in contradiction with our assumption. □

We can use the above result to study the structure of Σ near a point $x_0 \in \Sigma\setminus\Gamma$.

Theorem 3.3. *Let the point $x_0 \in \mathcal{R}\setminus\mathcal{K}$ be singular and not conjugate. Then, in a neighbourhood of x_0, the singular points of V are contained in a finite union of $(n-1)$-dimensional manifolds of class C^k.*

Proof. Since x_0 is not conjugate, by the previous theorem

$$\mathcal{F}(x_0) = \{\xi_1,\ldots,\xi_N\}.$$

Then,

$$D^*V(x_0) = \{q_1, \dots, q_N\},$$

where, denoted by τ_j the exit time of the optimal trajectory starting at x_0 and arriving in ξ_j, q_j is $P(\xi_j, \tau_j)$. We want to construct in a neighbourhood of x_0 smooth functions $V_1, \dots, V_N$ such that

$$V(x) = \min\{V_1(x), \dots, V_N(x)\}$$

and $DV_j(x_0) = q_j$. Let $W_1, \dots, W_N$ be neighbourhoods of $\xi_1, \dots, \xi_N$ with $\overline{W}_i \cap \overline{W}_j = \emptyset$ if $i \neq j$. Set $\mathcal{K}_j = \overline{W_j \cap \mathcal{K}}$ for $j = 1, \dots, N$. The set-valued function $\mathcal{F}(\cdot)$ is upper semicontinuous and so there exists $r > 0$ such that

$$\text{if} \quad x \in B_r(x_0), \quad \text{then} \quad \mathcal{F}(x) \subset \mathcal{K}_1 \cup \cdots \cup \mathcal{K}_N.$$

Let V_j be the value function of the exit time problem with target $\mathcal{K}_j$. Then,

$$V(x) = \min_{j=1,\dots,N} V_j(x), \quad \forall x \in B_r(x_0).$$

Moreover, $D^*V_j(x_0) = \{DV_j(x_0)\} = \{q_j\}$, since from x_0 starts only one optimal trajectory considering $\mathcal{K}_j$ as target. Since x_0 is regular and not conjugate for V_j, $j = 1, \dots, N$, the functions V_j are of class C^k in a neighbourhood of x_0. This implies that in a neighbourhood of x_0, the points where V is not differentiable are contained in the union of sets of type $\{x\colon V_i(x) = V_j(x)\}$, for $i \neq j$. Since $DV_i(x_0) \neq DV_j(x_0)$, these sets are locally $(n-1)$-dimensional manifolds of class C^k. $\square$

4. Rectifiability of $\overline{\Sigma}$

We know, by Corollary 3.2, that the inclusion $\overline{\Sigma} \subset \Sigma \cup \Gamma$ holds. Then, if we prove the $\mathcal{H}^{n-1}$-rectifiability of Γ we can conclude, recalling Theorem 1.2, that $\overline{\Sigma}$ is $\mathcal{H}^{n-1}$-rectifiable.

In this section we assume the additional hypothesis:

(H3) The Hamiltonian $H(x, p)$ is strongly convex in p, that is

$$H_{pp}(x, p) > 0, \quad \forall (x, p) \in \mathbf{R}^n \times \big(\mathbf{R}^n \setminus \{0\}\big).$$

Remark 4.1. In the case of the problem considered in Example 2.1 the value function V satisfies the Hamilton–Jacobi equation $H(x, DV(x)) = 0$, where the Hamiltonian

$$H(x, p) = r\left|\sigma^T(x)p\right| - L(x)$$

is convex but not strongly convex in p. However, we can observe that the Hamiltonian

$$\tilde{H}(x, p) = \frac{1}{2}\left(r^2\left|\sigma^T(x)p\right|^2 - L^2(x)\right)$$

has the same viscosity solution V and it is strongly convex in the variable p. Moreover, it is easy to see that the control problem associated to $\tilde{H}$ corresponds to the choices

$$\tilde{f}(x,u) = r\sigma(x)u, \qquad U = \mathbf{R}^n, \qquad \tilde{L}(x) = \frac{L^2(x)}{2} + \frac{|u|^2}{2}.$$

For a problem of this form the maximum principle is known. Therefore, since the value function V is semiconcave, also for the new control problem we have the Hamiltonian formulation of the maximum principle. Thus, in our study, we can replace H with $\tilde{H}$. If we compute the derivatives of $\tilde{H}$ along the characteristics, we obtain

$$\tilde{H}_p(x,p) = r\left|\sigma^T(x)p\right| H_p(x,p),$$
$$\tilde{H}_x(x,p) = r\left|\sigma^T(x)p\right| H_x(x,p).$$

So, the set Γ of the conjugate points does not change replacing H with $\tilde{H}$. We can proceed analogously in the case of Example 2.2.

We recall the following Sard-type result (see [18]).

Theorem 4.1. *Let $A \subset \mathbf{R}^n$ be an open set and let $F : A \subset \mathbf{R}^n \to \mathbf{R}^m$ be a map of class C^k, for $k \geqslant 1$. For any $j \in \{0, 1, \ldots, m-1\}$ set*

$$A_j = \left\{x \in \mathbf{R}^n \colon \operatorname{rk} DF(x) \leqslant j\right\},$$

where $\operatorname{rk} DF$ denotes the rank of DF. Then,

$$\mathcal{H}^{j+(n-j)/k}(F(A_j)) = 0.$$

We set

$$G = \left\{(\xi,t)\colon \det Y_{\xi,t}(\xi,t) = 0\right\}.$$

By definition $\Gamma \subset Y(G)$. We can use this fact to obtain upper bounds on the size of Γ.

Applying Theorem 4.1 to Y we obtain

$$\mathcal{H}^{n-1+1/k}(\Gamma) = 0.$$

However, that does not imply the rectifiability of Γ even if the data of our problem are C^∞.

To prove the rectifiability of Γ we need some preliminary results. First we recall some elementary facts of linear algebra.

Let $A = (a_{ij})$ be a $n \times n$ matrix. We denote by A^+ the transpose of the matrix of the cofactors of A; that is, $AA^+ = A^+A = (\det A)I$. Then the following properties hold (see [9]):

$$\operatorname{rk} A = n-1 \quad \Rightarrow \quad \operatorname{rk} A^+ = 1, \tag{32}$$

$$\operatorname{rk} A < n-1 \quad \Rightarrow \quad \operatorname{rk} A^+ = 0. \tag{33}$$

In addition, if $A(\cdot)$ is time-dependent and of class C^1, then

$$\frac{d}{dt}\det A(t) = \operatorname{tr}\bigl(A'(t)A^+(t)\bigr). \tag{34}$$

We begin with some preliminary results.

Lemma 4.1. *For all $\xi \in \partial\mathcal{K}$ let $(Y(\xi,\cdot), P(\xi,\cdot))$ be the unique solution of system* (18). *Then*

$$Y^T_{\xi,t}(\xi,t)P_{\xi,t}(\xi,t) = P^T_{\xi,t}(\xi,t)Y_{\xi,t}(\xi,t), \quad t \geqslant 0, \tag{35}$$

where the Jacobians are understood as in Remark 3.2.

Proof. Let $\xi_0 \in \partial\mathcal{K}$ be fixed. First we prove the equality (35) for $t = 0$. For simplicity, let us denote by

$$\mathcal{M}(t) = (m_{ij}(t))_{i,j}$$

the matrix $Y^T_{\xi,t}(\xi_0,t)P_{\xi,t}(\xi_0,t)$. Let $\eta_0 \in \mathbf{R}^{n-1}$ and let φ be the local parametrization of $\partial\mathcal{K}$,

$$\varphi : V_{\eta_0} \subset \mathbf{R}^{n-1} \to I_{\xi_0} \subset \partial\mathcal{K},$$

such that $\xi_0 = \varphi(\eta_0)$. Therefore we can rewrite the initial conditions

$$\begin{cases} Y(\xi_0,0) = \xi_0 = \varphi(\eta_0), \\ P(\xi_0,0) = p_0 = Dg(\varphi(\eta_0)) + \mu(\varphi(\eta_0))\nu(\varphi(\eta_0)), \end{cases} \tag{36}$$

where we recall that, for $\xi \in \partial\mathcal{K}$, $\mu(\xi)$ is the unique positive constant such that

$$H\bigl(\xi, Dg(\xi) + \mu(\xi)\nu(\xi)\bigr) = 0. \tag{37}$$

It is easy to see that for $j = 1,\dots,n-1$,

$$m_{nj}(0) = H_p(\xi_0,p_0)\frac{\partial p_0}{\partial\eta_j}, \qquad m_{jn}(0) = -H_x(\xi_0,p_0)\frac{\partial\varphi}{\partial\eta_j}(\eta_0).$$

Recalling relation (37), by our regularity assumptions and by Proposition 3.2 we can derive the identity

$$H\bigl(\varphi(\eta), P(\varphi(\eta),0)\bigr) = 0,$$

with respect to the variable η, and therefore we obtain

$$m_{nj}(0) = m_{jn}(0), \quad \forall j = 1,\dots,n-1.$$

In the case $i \neq j$ and $i, j \neq n$,

$$m_{ij}(0) = \frac{\partial Y}{\partial\eta_i}(\xi_0,0)\cdot\frac{\partial P}{\partial\eta_j}(\xi_0,0).$$

Calculating the derivatives we obtain

$$
\begin{aligned}
m_{ij}(0) = {} & \sum_{k,l=1}^{n} \frac{\partial^2 g}{\partial x_k \partial x_l}(\varphi(\eta_0)) \frac{\partial \varphi_k}{\partial \eta_j}(\eta_0) \frac{\partial \varphi_l}{\partial \eta_i}(\eta_0) \\
& + \sum_{k,l=1}^{n} \frac{\partial \varphi_l}{\partial \eta_i}(\eta_0) \frac{\partial \varphi_k}{\partial \eta_j}(\eta_0) \frac{\partial \mu}{\partial x_k}(\varphi(\eta_0)) \nu_l(\varphi(\eta_0)) \\
& + \mu(\varphi(\eta_0)) \frac{\partial \varphi}{\partial \eta_i}(\eta_0) \cdot \frac{\partial \nu}{\partial \eta_j}(\varphi(\eta_0)). \qquad (38)
\end{aligned}
$$

We note that the first term in the right-hand side of (38) is symmetric with respect to the indices i and j. Now, we observe that

$$
\frac{\partial \varphi}{\partial \eta_i}(\eta) \cdot \nu(\varphi(\eta)) = 0, \quad \text{for } \eta \in V_{\eta_0},\ i = 1, \dots, n-1. \qquad (39)
$$

So, the second term in the right-hand side of (38) is null. It remains to prove the equality

$$
\frac{\partial \varphi}{\partial \eta_i}(\eta) \frac{\partial \nu}{\partial \eta_j}(\varphi(\eta)) = \frac{\partial \varphi}{\partial \eta_j}(\eta) \frac{\partial \nu}{\partial \eta_i}(\varphi(\eta)), \quad \text{for } \eta \in V_{\eta_0}. \qquad (40)
$$

By (39) we have

$$
\begin{aligned}
\frac{\partial \varphi}{\partial \eta_i}(\eta) \frac{\partial \nu}{\partial \eta_j}(\varphi(\eta)) &= \frac{\partial}{\partial \eta_j}\left(\nu(\varphi(\eta)) \cdot \frac{\partial \varphi}{\partial \eta_i}(\eta) \right) - \nu(\varphi(\eta)) \cdot \frac{\partial^2 \varphi}{\partial \eta_i \partial \eta_j}(\eta) \\
&= -\nu(\varphi(\eta)) \cdot \frac{\partial^2 \varphi}{\partial \eta_i \partial \eta_j}(\eta),
\end{aligned}
$$

and this gives (40). So, $m_{ij}(0) = m_{ji}(0)$, also in the case $i \neq j$ and $i, j \neq n$. Then the matrix $\mathcal{M}(0)$ is symmetric. This proves (35) for $t = 0$. Now, by a simple computation, from the linear system (25) we obtain

$$
\begin{aligned}
\mathcal{M}'(t) = {} & P_{\xi,t}^T(\xi_0, t) H_{pp}\big(Y(\xi_0, t), P(\xi_0, t)\big) P_{\xi,t}(\xi_0, t) \\
& - Y_{\xi,t}^T(\xi_0, t) H_{xx}\big(Y(\xi_0, t), P(\xi_0, t)\big) Y_{\xi,t}(\xi_0, t),
\end{aligned}
$$

that is symmetric.

Therefore

$$
\mathcal{M}'(t) = \mathcal{M}^{T\,\prime}(t),
$$

and then, since $\mathcal{M}(0) = \mathcal{M}^T(0)$, the equality (35) holds for all $t \geqslant 0$. $\square$

Lemma 4.2. *Given* $\xi_0 \in \partial\mathcal{K}$ *and* $t_0 > 0$*, let* $\operatorname{rk} Y_{\xi,t}(\xi_0, t_0) = n - 1$ *and let* θ *be a generator of* $\operatorname{Ker} Y_{\xi,t}(\xi_0, t_0)$*. Then,* $P_{\xi,t}(\xi_0, t_0)\theta$ *generates the orthogonal complement of* $\operatorname{Im} Y_{\xi,t}(\xi_0, t_0)$ *and*

$$
Y_{\xi,t}^{+}(\xi_0, t_0) = c\theta \otimes P_{\xi,t}(\xi_0, t_0)\theta \qquad (41)
$$

for some $c \in \mathbf{R} \setminus \{0\}$.

Proof. Let $\eta_0 \in \mathbf{R}^{n-1}$ and let φ be the local parametrization of $\partial\mathcal{K}$,

$$\varphi : V_{\eta_0} \subset \mathbf{R}^{n-1} \to I_{\xi_0} \subset \partial\mathcal{K},$$

such that $\varphi(\eta_0) = \xi_0$. Observing that

$$Y_{\xi,t}(\xi_0, 0) = \left(\frac{\partial\varphi}{\partial\eta}(\eta_0), H_p(\xi_0, p_0)\right),$$

where, for simplicity, $p_0 = Dg(\xi_0) + \mu(\xi_0)\nu(\xi_0)$, it is easy to compute that $\det Y_{\xi,t}(\xi_0, 0) = cH_p(\xi_0, p_0) \cdot \nu(\xi_0)$, for a real constant $c \neq 0$. Then, by Proposition 3.1 we have that

$$\det Y_{\xi,t}(\xi_0, 0) \neq 0.$$

Therefore,

$$\operatorname{rank}\begin{pmatrix} Y_{\xi,t}(\xi_0, 0) \\ P_{\xi,t}(\xi_0, 0) \end{pmatrix} = n,$$

and by well-known properties of linear systems (see [21])

$$\operatorname{rank}\begin{pmatrix} Y_{\xi,t}(\xi_0, t) \\ P_{\xi,t}(\xi_0, t) \end{pmatrix} = \operatorname{rank}\begin{pmatrix} Y_{\xi,t}(\xi_0, 0) \\ P_{\xi,t}(\xi_0, 0) \end{pmatrix} = n, \quad t \in [0, +\infty).$$

Since $Y_{\xi,t}(\xi_0, t_0)\theta = 0$, we have $P_{\xi,t}(\xi_0, t_0)\theta \neq 0$.

Given any $w \in \mathbf{R}^n$, we have

$$\begin{aligned}\langle P_{\xi,t}(\xi_0, t_0)\theta, Y_{\xi,t}(\xi_0, t_0)w\rangle &= \langle Y^T_{\xi,t}(\xi_0, t_0)P_{\xi,t}(\xi_0, t_0)\theta, w\rangle \\ &= \langle P^T_{\xi,t}(\xi_0, t_0)Y_{\xi,t}(\xi_0, t_0)\theta, w\rangle = 0.\end{aligned}$$

Since w is arbitrary, this proves that $P_{\xi,t}(\xi_0, t_0)\theta$ generates the orthogonal complement of $\operatorname{Im} Y_{\xi,t}(\xi_0, t_0)$. Now we want to prove (41). We observe that, by (32), $Y^+_{\xi,t}(\xi_0, t_0)$ has rank one, and so $Y^+_{\xi,t}(\xi_0, t_0) = v_1 \otimes v_2$ for some nonzero vectors v_1, v_2.

Moreover,

$$v_1\mathbf{R} = \operatorname{Im} Y^+_{\xi,t}(\xi_0, t_0) = \operatorname{Ker} Y_{\xi,t}(\xi_0, t_0).$$

Indeed, by definition of $Y^+_{\xi,t}$ it follows that $\operatorname{Im} Y^+_{\xi,t}(\xi_0, t_0) \subset \operatorname{Ker} Y_{\xi,t}(\xi_0, t_0)$; but since both $\operatorname{Im} Y^+_{\xi,t}(\xi_0, t_0)$ and $\operatorname{Ker} Y_{\xi,t}(\xi_0, t_0)$ are one-dimensional spaces, they are the same space. Analogously, we have

$$\{v_2\}^\perp = \operatorname{Ker} Y^+_{\xi,t}(\xi_0, t_0) = \operatorname{Im} Y_{\xi,t}(\xi_0, t_0).$$

Therefore v_1 and v_2 have to be parallel, respectively, to θ and $P_{\xi,t}(\xi_0, t_0)\theta$. This ends the proof. $\square$

Lemma 4.3. *Let* $(\xi_0, t_0) \in G$. *Then,*

$$\frac{d}{dt}\det Y_{\xi,t}(\xi_0, t)\Big|_{t=t_0} \neq 0 \quad \Leftrightarrow \quad \operatorname{rk} Y_{\xi,t}(\xi_0, t_0) = n - 1.$$

Proof. By (34) we have

$$\begin{aligned}
&\frac{d}{dt}\det Y_{\xi,t}(\xi_0,t)\Big|_{t=t_0}\\
&\quad=\operatorname{tr}\bigl(Y'_{\xi,t}(\xi_0,t_0)Y^+_{\xi,t}(\xi_0,t_0)\bigr)\\
&\quad=\operatorname{tr}\bigl(H_{px}Y_{\xi,t}(\xi_0,t_0)Y^+_{\xi,t}(\xi_0,t_0)\bigr)+\operatorname{tr}\bigl(H_{pp}P_{\xi,t}(\xi_0,t_0)Y^+_{\xi,t}(\xi_0,t_0)\bigr)\\
&\quad=\det Y_{\xi,t}(\xi_0,t_0)\operatorname{tr}(H_{px})+\operatorname{tr}\bigl(H_{pp}P_{\xi,t}(\xi_0,t_0)Y^+_{\xi,t}(\xi_0,t_0)\bigr)\\
&\quad=\operatorname{tr}\bigl(H_{pp}P_{\xi,t}(\xi_0,t_0)Y^+_{\xi,t}(\xi_0,t_0)\bigr),
\end{aligned}$$

where, for simplicity, we omitted the dependence of H on Y, P.

Now, we suppose that $\operatorname{rk} Y_{\xi,t}(\xi_0,t_0)=n-1$. If θ is such that $\operatorname{Ker} Y_{\xi,t}(\xi_0,t_0)=\theta\mathbf{R}$, we have, by the previous lemma,

$$\begin{aligned}
\frac{d}{dt}\det Y_{\xi,t}(\xi_0,t)\Big|_{t=t_0}&=\operatorname{tr}\bigl(cH_{pp}P_{\xi,t}(\xi_0,t_0)\theta\otimes P_{\xi,t}(\xi_0,t_0)\theta\bigr)\\
&=cH_{pp}P_{\xi,t}(\xi_0,t_0)\theta\cdot P_{\xi,t}(\xi_0,t_0)\theta\neq 0
\end{aligned}$$

since H_{pp} is positive definite and $P_{\xi,t}(\xi_0,t_0)\theta\neq 0$.

On the other hand, if $\operatorname{rk} Y_{\xi,t}(\xi_0,t_0)<n-1$, then $Y^+_{\xi,t}(\xi_0,t_0)=0$ by (33) and therefore

$$\frac{d}{dt}\det Y_{\xi,t}(\xi_0,t)\Big|_{t=t_0}=0.$$

This ends the proof. □

We are ready to give the rectifiability theorem.

Theorem 4.2. *The set Γ of the conjugate points of the value function V is $\mathcal{H}^{n-1}$-rectifiable.*

Proof. Let us define

$$G'=\bigl\{(\xi,t):\ \operatorname{rk} Y_{\xi,t}(\xi,t)=n-1\bigr\},\tag{42}$$

$$G''=\bigl\{(\xi,t):\ \operatorname{rk} Y_{\xi,t}(\xi,t)<n-1\bigr\}.\tag{43}$$

Then, by Lemma 4.3 and by the implicit function theorem, we can conclude that G' is $\mathcal{H}^{n-1}$-rectifiable. Since Y is of class C^k also $Y(G')$ is $\mathcal{H}^{n-1}$-rectifiable.

On the other hand, by Theorem 4.1,

$$\mathcal{H}^{n-1}\bigl(Y(G'')\bigr)\leqslant\mathcal{H}^{n-2+2/k}\bigl(Y(G'')\bigr)=0.$$

Since $\Gamma\subset Y(G'\cup G'')$, the result is proved. □

As announced, by Theorem 4.2 it follows

Corollary 4.1. *The closure $\overline{\Sigma}$ of the singular set is $\mathcal{H}^{n-1}$-rectifiable.*

Remark 4.2. As an application of the rectifiability of $\overline{\Sigma}$ it follows that $DV \in SBV_{\mathrm{loc}}(\mathcal{R} \setminus \mathcal{K})$, where *SBV* is the class of special functions of bounded variation introduced by De Giorgi and Ambrosio (see [5,6]).

For the proof we refer to [9] where a problem of calculus of variations is studied. The proof in our case is entirely analogous.

5. Further Hausdorff estimates

We can give a better Hausdorff estimate for the set of the regular and conjugate points $\Gamma \setminus \Sigma$. We need some preliminary results.

Remark 5.1. As a consequence of the implicit function theorem we have that the set G' defined in (42) is locally a graph. That is, for any $(\xi_0, t_0) \in G'$ there exist constants $r, \varepsilon > 0$ and a function

$$\phi : B_r(\xi_0) \cap \partial\mathcal{K} \to (t_0 - \varepsilon, t_0 + \varepsilon)$$

of class C^{k-1} such that

$$\det Y_{\xi,t}(\xi, t) = 0 \quad \Leftrightarrow \quad t = \phi(\xi)$$

for any $(\xi, t) \in (B_r(\xi_0) \cap \partial\mathcal{K}) \times (t_0 - \varepsilon, t_0 + \varepsilon)$. Choosing r small enough we can suppose that

$$\operatorname{rk} Y_{\xi,t}\big(\xi, \phi(\xi)\big) = n - 1 \quad \text{for any } \xi \in B_r(\xi_0) \cap \partial\mathcal{K}.$$

So, there exists a vector field

$$\theta : B_r(\xi_0) \cap \partial\mathcal{K} \to \mathbf{R}^n$$

of class C^{k-1} such that for any $\xi \in B_r(\xi_0) \cap \partial\mathcal{K}$, $\theta(\xi)$ is a generator of $\operatorname{Ker} Y_{\xi,t}(\xi, \phi(\xi))$. Indeed, it suffices take for θ the exterior product of $n-1$ linearly independent rows of $Y_{\xi,t}(\xi, \phi(\xi))$.

We want to study the relationship between the vector functions $D\phi$ and θ at a regular conjugate point. Let us denote by θ_i, $i = 1, \ldots, n$, the scalar components of θ and by θ' the vector field $(\theta_1, \ldots, \theta_{n-1})$. We give the following lemma.

Lemma 5.1. *Let $(\xi_0, t_0) \in G'$ and let the constants r, ε and the vector fields ϕ, θ be defined as in Remark* 5.1. *Then, for any $\xi \in B_r(\xi_0) \cap \partial\mathcal{K}$,*

$$D\phi(\xi) \cdot \theta'(\xi) = \theta_n(\xi) \quad \Leftrightarrow \quad \frac{\partial^2 Y}{\partial\theta(\xi)^2}\big(\xi, \phi(\xi)\big) \cdot \big(P_\xi\big(\xi, \phi(\xi)\big)\theta(\xi)\big) = 0.$$

Proof. By definition of $\theta(\xi)$, for any $\xi \in B_r(\xi_0) \cap \partial\mathcal{K}$ and $i = 1, \ldots, n$, we have

$$\sum_{j=1}^{n-1} \frac{\partial Y_i}{\partial \eta_j}\big(\xi, \phi(\xi)\big)\theta_j(\xi) + \frac{\partial Y_i}{\partial t}\big(\xi, \phi(\xi)\big)\theta_n(\xi) = 0. \tag{44}$$

Denote for simplicity $q(\xi) = P_{\xi,t}(\xi, \phi(\xi))\theta(\xi)$. By Lemma 4.2, $q(\xi)$ generates the orthogonal complement of $\operatorname{Im} Y_{\xi,t}(\xi, \phi(\xi))$. Differentiating (44) with respect to ξ_k and recalling (24) we obtain

$$\sum_{h=1}^{n} \frac{\partial^2 H}{\partial p_i \partial p_h} q_h \frac{\partial \phi}{\partial \eta_k} + \sum_{j=1}^{n-1} \frac{\partial^2 Y_i}{\partial \eta_j \partial \eta_k} \theta_j + \frac{\partial^2 Y_i}{\partial t \partial \eta_k} \theta_n + \sum_{j=1}^{n-1} \frac{\partial Y_i}{\partial \eta_j} \frac{\partial \theta_j}{\partial \eta_k} + \frac{\partial Y_i}{\partial t} \frac{\partial \theta_n}{\partial \eta_k} = 0 \tag{45}$$

for $i = 1, \ldots, n$, and $k = 1, \ldots, n-1$. Multiplying (45) by $q_i \theta_k$ and summing over i, k, we obtain

$$\sum_{i,h=1}^{n} \frac{\partial^2 H}{\partial p_i \partial p_h} q_h q_i \sum_{k=1}^{n-1} \frac{\partial \phi}{\partial \eta_k} \theta_k + \sum_{i=1}^{n} \sum_{j,k=1}^{n-1} \frac{\partial^2 Y_i}{\partial \eta_j \partial \eta_k} \theta_j \theta_k q_i + \sum_{i=1}^{n} \sum_{k=1}^{n-1} \frac{\partial^2 Y_i}{\partial t \partial \eta_k} \theta_n \theta_k q_i = 0, \tag{46}$$

where we used the fact that q is orthogonal to $\operatorname{Im} Y_{\xi,t}$. Equality (46) can be rewritten as

$$(H_{pp} q \cdot q)(D\phi \cdot \theta' - \theta_n) + \frac{\partial^2 Y}{\partial \theta^2} q = 0.$$

Since H_{pp} is strictly positive and $q \neq 0$, the result is proved. □

Now, we recall a result about the local invertibility of functions. For the proof we refer, e.g., to [4].

Lemma 5.2. *Let $F : \mathbf{R}^m \to \mathbf{R}^m$ be a map of class C^2, let $\bar{z}$ be a point at which* $\operatorname{rk} DF(\bar{z}) = m - 1$, *and set $\bar{x} = F(\bar{z})$. Let us consider a generator θ of* $\operatorname{Ker} DF(\bar{z})$ *and a nonzero vector w orthogonal to* $\operatorname{Im} DF(\bar{z})$. *Suppose that*

$$\frac{\partial^2 F}{\partial \theta^2}(\bar{z}) \cdot w > 0.$$

Then, there exist positive constants ρ, σ such that the equation

$$F(z) = \bar{x} + sw, \quad z \in B_\rho(\bar{z}),$$

has two solutions if $0 < s < \sigma$, *and no solution if* $-\sigma < s < 0$.

Using Lemmas 5.1 and 5.2 we give the following result.

Proposition 5.1. *Given $(\xi_0, t_0) \in G'$, let r, ε and ϕ, θ be defined as in Remark* 5.1. *If $\bar{\xi} \in B_r(\xi_0) \cap \partial\mathcal{K}$ is such that $Y(\bar{\xi}, \phi(\bar{\xi})) \notin \Sigma$, then*

$$D\phi(\bar{\xi}) \cdot \theta'(\bar{\xi}) = \theta_n(\bar{\xi}).$$

Proof. Set $\bar{t} = \phi(\bar{\xi})$ and $\bar{x} = Y(\bar{\xi}, \bar{t})$. Arguing by contradiction we suppose that $D\phi(\bar{\xi}) \cdot \theta'(\bar{\xi}) \neq \theta_n(\bar{\xi})$. Then, by Lemma 5.1,

$$\frac{\partial^2 Y}{\partial \theta^2}(\bar{\xi}, \bar{t}) \cdot w \neq 0, \quad \text{if } w = \pm P_{\xi,t}(\bar{\xi}, \bar{t})\theta(\bar{\xi}).$$

Then we can chose the sign of w to have

$$\frac{\partial^2 Y}{\partial \theta^2}(\bar{\xi}, \bar{t}) \cdot w > 0.$$

We note that θ belongs to $\operatorname{Ker} Y_{\xi,t}(\bar{\xi}, \bar{t})$ and that w is a nonzero vector generating the orthogonal of $\operatorname{Im} Y_{\xi,t}(\bar{\xi}, \bar{t})$. Therefore, the map $Y(\cdot\,,\cdot): \partial\mathcal{K} \times [0, +\infty) \to \mathbf{R}^n$ satisfies all the assumptions of Lemma 5.2. Define $x_k = \bar{x} - k^{-1} w$ and let y_k be the optimal trajectories starting at x_k with exit time τ_k. Consider the sequence $\xi_k = y_k(\tau_k)$. Then $x_k = Y(\xi_k, \tau_k)$. So, by Lemma 5.2, $(\xi_k, \tau_k) \notin B_\rho(\bar{\xi}, \bar{t})$, for large k. But $\{\xi_k\}_{k\in\mathbf{N}}$ is bounded; then up to a subsequence $\xi_k \to \tilde{\xi} \neq \bar{\xi}$. Recalling Proposition 3.3, this means that there exist an optimal characteristic $Y(\tilde{\xi}, \cdot)$ different from $Y(\bar{\xi}, \cdot)$ and a time $\tilde{\tau}$ such that

$$Y(\tilde{\xi}, \tilde{\tau}) = Y(\bar{\xi}, \bar{t}) = \bar{x}.$$

This is in contradiction with the fact that $\bar{x}$ is a regular point. □

Lemma 5.3. *Let* $(\xi_0, t_0) \in G'$, *and let the constants* r, ε *and the vector fields* ϕ, θ *be defined as in Remark* 5.1. *If for* $\xi \in B_r(\xi_0)$, $\operatorname{rk} Y_\xi(\xi, \phi(\xi)) = n - 2$, *then* $\theta_n(\xi) = 0$ *and* $\theta'(\xi) = (\theta_1(\xi), \ldots, \theta_{n-1}(\xi))$ *generates* $\operatorname{Ker} Y_\xi(\xi, \phi(\xi))$.

Proof. If $\operatorname{rk} Y_\xi(\xi, \phi(\xi)) = n - 2$, then there exists $w \in \mathbf{R}^{n-1}$, $w \neq 0$, such that $Y_\xi(\xi, \phi(\xi)) \cdot w = 0$. The vector $(w, 0) \in \mathbf{R}^n$ is a generator of $\operatorname{Ker} Y_{\xi,t}(\xi, \phi(\xi))$. Since $\dim \operatorname{Ker} Y_{\xi,t}(\xi, \phi(\xi)) = 1$, this implies that $\theta(\xi)$ and $(w, 0)$ are linearly dependent. So, $\theta_n(\xi) = 0$ and $\theta'(\xi)$ generates $\operatorname{Ker} Y_\xi(\xi, \phi(\xi))$. □

Remark 5.2. By Proposition 5.1 and Lemma 5.3 it follows that if $Y(\xi, \phi(\xi)) \in \Gamma \setminus \Sigma$, then $\theta'(\xi) \neq 0$.

Now we can give the announced Hausdorff estimate about the set $\Gamma \setminus \Sigma$.

Theorem 5.1. *If* $\Gamma \setminus \Sigma$ *is the set of the conjugate regular points of the value function* V, *then*

$$\mathcal{H}^{n-2+2/k}(\Gamma \setminus \Sigma) = 0.$$

Proof. Let $(\xi_0, t_0) \in G'$. Then by Remark 5.1 there exist $r, \varepsilon > 0$ such that $[(B_r(\xi_0) \cap \partial\mathcal{K}) \times (t_0 - \varepsilon, t_0 + \varepsilon)] \cap G$ is the graph of a function $\phi : B_r(\xi_0) \cap \partial\mathcal{K} \to (t_0 - \varepsilon, t_0 + \varepsilon)$. We define

$$\tilde{Y}(\xi) = Y\big(\xi, \phi(\xi)\big), \quad \xi \in B_r(\xi_0) \cap \partial\mathcal{K}.$$

Consider $\xi \in B_r(\xi_0)$ with $\tilde{Y}(\xi) \notin \Sigma$. There are two eventualities: $\operatorname{rk} Y_\xi(\xi, \phi(\xi)) = n-2$ or $\operatorname{rk} Y_\xi(\xi, \phi(\xi)) = n-1$. If $\operatorname{rk} Y_\xi(\xi, \phi(\xi)) = n-2$, then by Proposition 5.1 and Lemma 5.3, there exists a nonzero vector $\theta' \in \mathbf{R}^{n-1}$ such that $Y_\xi(\xi, \phi(\xi))\theta' = D\phi(\xi) \cdot \theta' = 0$. So, for such a ξ,

$$\operatorname{rk} D\tilde{Y}(\xi) = \operatorname{rk}\big[Y_\xi\big(\xi, \phi(\xi)\big) + Y_t\big(\xi, \phi(\xi)\big) \otimes D\phi(\xi)\big] \leqslant n-2.$$

Otherwise, if $\operatorname{rk} Y_\xi(\xi, \phi(\xi)) = n-1$, using Proposition 5.1 it is easy to compute that

$$D\tilde{Y}(\xi)\theta'(\xi) = Y_{\xi,t}\big(\xi, \phi(\xi)\big)\theta(\xi) = 0.$$

So, also in this case, $\operatorname{rk} D\tilde{Y}(\xi) \leqslant n-2$. Therefore, by Theorem 4.1 we obtain that $\mathcal{H}^{n-2+1/k}(Y(G') \cap (\Gamma \setminus \Sigma)) = 0$. On the other hand, we already know that $\mathcal{H}^{n-2+2/k}(Y(G'')) = 0$. Thus, the conclusion immediately follows, being $k \geqslant 2$. $\square$

Remark 5.3. We note that if our data are of class C^∞, then

$$\mathcal{H} - \dim(\Gamma \setminus \Sigma) \leqslant n-2.$$

6. Conjugate points of the minimum time function

In the case of the minimum time function we can give an equivalent characterization of conjugate points considering only the spatial derivatives.

Theorem 6.1. *Consider the minimum time problem* ($L \equiv 1$ *and* $g \equiv 0$). *Let* $x_0 \in \Gamma$ *and let* $\xi_0 \in \mathcal{F}(x_0)$, $\xi_0 = y^{x_0,u}(T(x_0))$, *such that* $\det Y_{\xi,t}(\xi_0, T(x_0)) = 0$. *Then,*

$$\operatorname{rk} Y_\xi\big(\xi_0, T(x_0)\big) < n-1.$$

Proof. We argue by contradiction. Let us suppose that $\operatorname{rk} Y_\xi(\xi_0, T(x_0)) = n-1$. Then, the vectors $Y_{\eta_i}(\xi_0, T(x_0))$, $i = 1, \ldots, n-1$, are linearly independent. By continuity of $Y_\xi(\xi_0, \cdot)$, there exists a positive number δ such that the vectors $Y_{\eta_i}(\xi_0, t)$ are linearly independent for any time $t \in (T(x_0) - \delta, T(x_0))$. We need to distinguish two cases.

Case 1. There exists a sequence of nonconjugate points $x_k = Y(\xi_0, t_k)$ with $t_k \to T(x_0)$ as $k \to \infty$.

Since $Y(\xi, \cdot)$ is an optimal characteristic the points x_k are regular. By Remark 3.4 and definition of T,

$$DT(x_k) \cdot Y_t(\xi_0, t_k) = 1, \quad k \in \mathbf{N}. \tag{47}$$

So, $DT(x_k) \neq 0$. We claim that $DT(x_k)$ is parallel to $\nu(\xi_0, t_k)$, where $\nu(\xi_0, t)$ denotes the outward versor normal to the surface

$$S_t = \big\{Y(\xi, t)\colon\ \xi \in \partial\mathcal{K}\big\}$$

at the point $Y(\xi_0, t)$. In fact, the level surface of the minimum time $T(x) = t_k$ is regular in a neighbourhood of x_k and it coincides with S_t. Therefore, taking the limit for $k \to \infty$,

$$\lim_{k\to\infty} DT(x_k) \quad \text{and} \quad \nu\big(\xi_0, T(x_0)\big)$$

are parallel. Since x_0 is conjugate

$$\det Y_{\xi,t}\big(\xi_0, T(x_0)\big) = cY_t\big(\xi_0, T(x_0)\big) \cdot \nu\big(\xi_0, T(x_0)\big) = 0, \quad c \neq 0,$$

and so, letting $k \to \infty$ in (47),

$$P\big(\xi_0, T(x_0)\big) \cdot Y_t\big(\xi_0, T(x_0)\big) = 0.$$

But since $Y_t(\xi_0, t) = -f(Y(\xi_0, t), u^*(Y(\xi_0, t), P(\xi_0, t)))$, this contradicts the maximum principle which implies

$$-P(\xi_0, t) \cdot f\Big(Y(\xi_0, t), u^*\big(Y(\xi_0, t), P(\xi_0, t)\big)\Big) = 1.$$

Case 2. There exists a $\delta' \in (0, \delta)$ such that $Y(\xi_0, t)$ is conjugate for any $t \in (T(x_0) - \delta', T(x_0))$. In this case, since $\det Y_{\xi,t}(\xi_0, t)$ is constant, we have

$$\frac{d}{dt} \det Y_{\xi,t}(\xi_0, t) = 0, \quad \forall t \in \big(T(x_0) - \delta', T(x_0)\big).$$

Therefore, by Lemma 4.3, rk $Y_{\xi,t}(\xi_0, t) \neq n - 1$ in contrast with our assumption that $Y_{\eta_i}(\xi_0, t)$ are linearly independent. □

Remark 6.1. By Theorem 6.1 we immediately can give an equivalent definition of conjugate point in the case of the minimum time problem:

$x_0 \in \mathcal{R} \setminus \mathcal{K}$ is conjugate if there exists $\xi_0 \in \mathcal{F}(x_0)$, $\xi_0 = y^{x_0,u}(T(x_0))$, such that $\operatorname{rk} Y_\xi(\xi_0, T(x_0)) < n - 1$.

References

[1] P. Albano, P. Cannarsa, Singularities of semiconcave functions in Banach spaces, in: W.M. McEneaney, G.G. Yin, Q. Zhang (Eds.), Stochastic Analysis, Control, Optimization and Applications, Birkhäuser, Boston, 1999, pp. 171–190.

[2] P. Albano, P. Cannarsa, Structural properties of singularities of semiconcave functions, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 28 (1999) 719–740.

[3] G. Alberti, L. Ambrosio, P. Cannarsa, On the singularities of convex functions, Manuscripta Math. 76 (1992) 421–435.

[4] A. Ambrosetti, G. Prodi, A Primer on Nonlinear Analysis, Cambridge University Press, Cambridge, 1993.

[5] L. Ambrosio, Variational problems in *SBV* and image segmentation, Acta Appl. Math. 17 (1989) 1–40.

[6] L. Ambrosio, Existence theory for a new class of variational problems, Arch. Rational Mech. Anal. 111 (1990) 291–322.

[7] M. Bardi, I. Capuzzo Dolcetta, Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations, Birkhäuser, Boston, 1997.
[8] P. Cannarsa, H. Frankowska, C. Sinestrari, Optimality conditions and synthesis for the minimum time problem with a general target, Set-Valued Anal. 8 (2000) 127–148.
[9] P. Cannarsa, A. Mennucci, C. Sinestrari, Regularity results for solutions of a class of Hamilton–Jacobi equations, Arch. Rational Mech. Anal. 140 (1997) 197–223.
[10] P. Cannarsa, C. Pignotti, C. Sinestrari, Semiconcavity for optimal control problems with exit time, Discrete Contin. Dynam. Systems 6 (2000) 975–997.
[11] P. Cannarsa, C. Sinestrari, Convexity properties of the minimum time function, Calc. Variations 3 (1995) 273–298.
[12] P. Cannarsa, C. Sinestrari, On a class of nonlinear time optimal control problems, Discrete Contin. Dynam. Systems 1 (1995) 285–300.
[13] P. Cannarsa, H.M. Soner, On the singularities of the viscosity solutions to Hamilton–Jacobi–Bellman equations, Indiana Univ. Math. J. 36 (1987) 501–524.
[14] M.G. Crandall, L.C. Evans, P.L. Lions, Some properties of viscosity solutions of Hamilton–Jacobi equations, Trans. Amer. Math. Soc. 282 (1984) 487–502.
[15] M.G. Crandall, P.L. Lions, Viscosity solutions of Hamilton–Jacobi equations, Trans. Amer. Math. Soc. 277 (1983) 1–42.
[16] L.C. Evans, Partial Differential Equations, Graduate Studies in Mathematics, Vol. 19, American Mathematical Society, Providence, RI, 1998.
[17] H.O. Fattorini, Infinite Dimensional Optimization and Control Theory, Cambridge University Press, 1996.
[18] H. Federer, Geometric Measure Theory, Springer, Berlin, 1969.
[19] W.H. Fleming, The Cauchy problem for a nonlinear first order partial differential equation, J. Differential Equations 5 (1969) 515–530.
[20] W.H. Fleming, H.M. Soner, Controlled Markov Processes and Viscosity Solutions, Springer, New York, 1993.
[21] P. Hartman, Ordinary Differential Equations, Wiley, New York, 1964.
[22] P.L. Lions, Generalized Solutions of Hamilton–Jacobi Equations, Pitman, Boston, 1982.
[23] C. Mantegazza and A.C. Mennucci, Hamilton–Jacobi equations and distance functions on Riemannian manifolds, preprint No. 5, Scuola Normale Superiore, Pisa (1999).
[24] A.C. Mennucci, Regularity of solutions to Hamilton–Jacobi equations, preprint (2000).
[25] L. Zajíček, On the points of multiplicity of monotone operators, Comment. Math. Univ. Carolin. 19 (1978) 179–189.